*SOME NONPARAMETRIC SEQUENTIAL TESTS WITH POWER ONE**

BY D. A. DARLING AND HERBERT ROBBINS

UNIVERSITY OF CALIFORNIA, BERKELEY, AND UNIVERSITY OF MICHIGAN

Communicated by J. Neyman, August 28, 1968

1. *Introduction and Summary.*—We consider several one- and two-population nonparametric hypothesis-testing problems and derive sequential tests with power one and uniformly small error probability under the null hypothesis. Bounds for the expected sample size under the alternative hypothesis are given. These tests provide a first step toward the extension to the nonparametric case of the methods described in reference 3.

2. *The Tests.*—For any two distribution functions (d.f.'s) G,H denote

$$D^+(G,H) = \sup_{-\infty<t<\infty} (G(t) - H(t)), \qquad D(G,H) = \sup_{-\infty<t<\infty} |G(t) - H(t)|.$$

Let $0 < \epsilon < 1$ and a positive integer m be given, and let $f(x)$ denote any continuous, positive, nondecreasing function defined for $x \geqq m$ and such that (*i*) $f(x) \leqq x$; (*ii*) $f(x)/x$ is strictly decreasing to 0 as $x \to \infty$; (*iii*) $f(x)$ is concave; and

$$(iv) \sum_{n=m}^{\infty} \exp\left(\frac{-f^2(n)}{n+1}\right) \leqq \epsilon.$$

(Not all these restrictions are needed in some of the results that follow. Some examples of such functions are given in section 4.) For $0 < x \leqq f(m)/m$, define $g(x)$ to be the function inverse to $f(x)/x$.

Let $x_1,x_2,\ldots$ be independent, identically distributed (i.i.d.) with d.f. $F_x(t) = P(x_i \leqq t)$, and let $y_1,y_2,\ldots$ be i.i.d. with d.f. $F_y(t)$, the x's and y's being independent. Denote by $F_x^n(t) =$ (number of values $x_1,\ldots,x_n$ that are $\leqq t)/n$ the empirical d.f. of $x_1,\ldots,x_n$, and by $F_y^n(t)$ the empirical d.f. of $y_1,\ldots,y_n$ $(n \geqq 1)$.

(*a*) Consider the hypothesis

$$H_1{:}F_x(t) \leqq F_y(t) \text{ for every } -\infty < t < \infty.$$

THEOREM 1. *Define N as the smallest $n \geqq m$ such that $D^+(F_x^n,F_y^n) \geqq f(n)/n$ and reject H_1 iff $N < \infty$. Then*

$$\text{if } H_1 \text{ is true, } P(\text{reject } H_1) \leqq \epsilon; \tag{2.1}$$

$$\text{if } H_1 \text{ is false, } P(\text{reject } H_1) = 1. \tag{2.2}$$

If H_1 is false and $D^+(F_x,F_y) = d > 0$ with $d \leqq f(m)/m$, then

$$EN \leqq g\left(d - \frac{m}{g(d)}\right). \tag{2.3}$$

(*b*) Consider the hypothesis

$$H_2{:}F_x(t) = F_y(t) \quad \text{for every } -\infty < t < \infty.$$

THEOREM 2. *Define N as the smallest $n \geqq m$ such that $D(F_x^n,F_y^n) \geqq f(n)/n$, and reject H_2 iff $N < \infty$. Then*

$$\textit{if } H_2 \textit{ is true, } P(\textit{reject } H_2) \leqq 2\epsilon; \tag{2.4}$$

$$\textit{if } H_2 \textit{ is false, } P(\textit{reject } H_2) = 1. \tag{2.5}$$

If H_2 *is false and* $D(F_x,F_y) = d > 0$ *with* $d \leqq f(m)/m$, *then (2.3) holds.*

(*c*) Let F_0 be any specified d.f. and consider the hypothesis

$$H_3: F_x(t) \leqq F_0(t) \quad \text{for every } -\infty < t < \infty.$$

THEOREM 3. *Define* N *as the smallest* $n \geqq m$ *such that* $D^+(F_x{}^n,F_0) \geqq f(n)/n$, *and reject* H_3 *iff* $N < \infty$. *Then*

$$\textit{if } H_3 \textit{ is true, } P(\textit{reject } H_3) \leqq 2\sqrt{2}\epsilon; \tag{2.6}$$

$$\textit{if } H_3 \textit{ is false, } P(\textit{reject } H_3) = 1. \tag{2.7}$$

If H_3 *is false and* $D^+(F_x,F_0) = d > 0$ *with* $d \leqq f(m)/m$, *then (2.3) holds.*

(*d*) Let $\mathcal{G}$ be any class of d.f.'s closed under the D metric; e.g., $\mathcal{G}$ may consist of a single d.f. F_0 or of the set $N(\mu,\sigma^2)$, $-\infty < \mu < \infty$, $0 \leqq \sigma^2 < \infty$, etc. Consider the hypothesis

$$H_4: F_x \in \mathcal{G}.$$

Define $w_n = \inf\limits_{G \in \mathcal{G}} D(F_x{}^n,G)$.

THEOREM 4. *Define* N *as the smallest* $n \geqq m$ *such that* $w_n \geqq f(n)/n$ *and reject* H_4 *iff* $N < \infty$. *Then*

$$\textit{if } H_4 \textit{ is true, } P(\textit{reject } H_4) \leqq 4\sqrt{2}\epsilon; \tag{2.8}$$

$$\textit{if } H_4 \textit{ is false, } P(\textit{reject } H_4) = 1. \tag{2.9}$$

If H_4 *is false and* $\inf\limits_{G \in \mathcal{G}} D(F_x,G) = d > 0$ *with* $d \leqq f(m)/m$, *then (2.3) holds.*

3. *Proofs of Theorems.*—For integers $0 \leqq r \leqq n$, define

$$A(r,n) = \frac{(n!)^2}{(n-r)!(n+r)!}. \tag{3.1}$$

If F_x,F_y are arbitrary d.f.'s such that H_1 holds, then it may be shown that

$$P\left(D^+(F_x{}^n,F_y{}^n) \geq \frac{r}{n}\right) \leqq A(r,n). \tag{3.2}$$

(If $F_x \equiv F_y$ is continuous, the equality sign holds between the two terms of (3.2); this was shown by Gnedenko and Korolyuk[4] and follows from an elementary combinatorial argument. The validity of the inequality of (3.2) for arbitrary F_x,F_y satisfying H_1 is easily seen by representing the x's and y's as functions of independent random variables uniformly distributed on (0,1).) We shall now show that

$$A(r,n) \leqq \exp\left(\frac{-r^2}{n+1}\right). \tag{3.3}$$

It can be calculated (ref. 6, chap. 12, problem 31) that

$$-\log A(r,n) = \int_0^1 \frac{x^n(1-x^r)(1-x^{-r})}{(1-x)\log x}\,dx,$$

and by setting $x = e^{-t}$ and using the inequalities $\cosh y - 1 \geqq y^2/2$, $\sinh y/y \leqq e^y$, $y \geqq 0$, we obtain (3.3) from

$$-\log A(r,n) = \int_0^\infty e^{-(n+1/2)t}\,\frac{\cosh rt - 1}{\sinh (t/2)}\,\frac{dt}{t}$$

$$\geqq r^2 \int_0^\infty e^{-(n+1/2)t - t/2}\,dt = \frac{r^2}{n+1}.$$

Assuming now that H_1 is true, and denoting by x^* the smallest integer $\geqq x$, we have

$$P(N < \infty) \leqq \sum_{n=m}^{\infty} P(D^+(F_x{}^n, F_y{}^n) \geqq f(n)/n) \leqq \sum_{n=m}^{\infty} A(f(n)^*, n)$$

$$\leqq \sum_{n=m}^{\infty} \exp\,(-f^2(n)/(n+1)) \leqq \epsilon, \quad (3.4)$$

which proves (2.1). Equation (2.2) follows from the fact that if H_1 is false, then $D^+(F_x{}^n,F_y{}^n) \to D^+(F_x,F_y) = d > 0$ with probability 1 as $n \to \infty$, while by hypothesis $f(n)/n \to 0$, so that $P(N < \infty) = 1$.

To prove (2.3), assume for simplicity that t is a number such that $F_x(t) - F_y(t) = d$. The distribution of $n(F_x{}^n(t) - F_y{}^n(t))$ is that of $s_n = u_1 + \ldots + u_n$, where the u's are i.i.d. with $Eu_i = d$ and $|u_i| \leqq 1$. Let N' be the smallest $n \geqq m$ such that $s_n \geqq f(n)$; then $N \leqq N'$ and

$$s_{N'} \leqq \max\,(m, f(N') + 1) \leqq m + f(N').$$

It is easily proved that $EN' < \infty$ and by Wald's lemma on cumulative sums and Jensen's inequality, we obtain

$$dEN' \leqq m + f(EN'), \quad (3.5)$$

which gives an implicit upper bound for EN' and hence for EN. To obtain an explicit bound for EN', we observe that the equation

$$dx = m + f(x), \quad (x \geqq m) \quad (3.6)$$

has a unique solution x_0, and $EN \leqq EN' \leqq x_0$. If we write $x_0 = g(y_0)$, equation (3.6) becomes

$$y_0 = d - \frac{m}{g(y_0)}, \quad (3.7)$$

and iterating this we obtain

$$y_0 = d - \frac{m}{g\left(d - \dfrac{m}{g(y_0)}\right)}$$

so that

$$y_0 \geqq d - \frac{m}{g(d)}, \tag{3.8}$$

and hence $EN \leqq x_0 \leqq g\left(d - \dfrac{m}{g(d)}\right)$, which completes the proof of Theorem 1. Theorem 2 follows by similar reasoning.

The proof of Theorem 3 is based on the fact that if H_3 is true, F_0 being arbitrary, then

$$P\left(D^+(F_x{}^n, F_0) \geqq \frac{r}{n}\right) \leqq B(r,n) \equiv \sum_{j=r}^{n} \binom{n}{j} \frac{r}{n-j+r} \frac{(j-r)^j(n-j+r)^{n-j}}{n^n}. \tag{3.9}$$

(If $F_x \equiv F_0$, a continuous d.f., then equality holds in (3.9), as was shown independently by Smirnov,[5] and Birnbaum and Tingey.[1]) We shall show that $B(r,n) \leqq 2\sqrt{2}A(r,n)$, and the method of proof of Theorem 1 shall then yield the proof of Theorem 3.

Using the solution of the classical ballot problem, we establish the curious identity

$$A(r,n) = \sum_{j=r}^{n} \frac{r}{2(n-j)+r} \cdot \frac{\binom{2j-r}{j}\binom{2(n-j)+r}{n-j}}{\binom{2n}{n}}. \tag{3.10}$$

(In an election in which A and B each receive n votes, let E_j be the event that when A has j votes he is, for the last time, r votes ahead of B. The jth term in the sum of (3.10) is $P(E_j)$.) Denote by R_j the ratio of the jth term of (3.9) to that of (3.10). We complete the proof by showing that $R_j \leqq 2\sqrt{2}$. After some algebra we have

$$R_j = \frac{ac}{b(j-r,j)b(n-j+r,n-j)},$$

with

$$a = \frac{2(n-j)+r}{n-j+r} = 2 - \frac{r}{n-j+r} \leqq 2,$$

$$b(x,i) = \left(1 + \frac{1}{x}\right)\left(1 + \frac{2}{x}\right)\cdots\left(1 + \frac{i}{x}\right), \qquad c = \frac{(2n)!}{n!n^n}.$$

Now

$$\log b(x,i) \geqq \int_0^i \log\left(1 + \frac{y}{x}\right) dy = (x+i)\log\left(1 + \frac{i}{x}\right) - i,$$

$$b(x,i) \geqq \left(1 + \frac{i}{x}\right)^{x+i} e^{-i}, \quad b(j-r,j)b(n-j+r,n-j)$$

$$\geqq \left(1 + \frac{j}{j-r}\right)^{2j-r} \left(1 + \frac{n-j}{n-j+r}\right)^{2(n-j)+r} e^{-n}.$$

If we set $\alpha = 2j - r$, $\beta = 2(n - j) + r$, further algebra and the inequality $1 + x \leqq e^x$ yield

$$b(j - r,j)b(n - j + r,n - j) \geqq \left(1 - \frac{r}{\alpha}\right)^{-\alpha} \left(1 + \frac{r}{\beta}\right)^{-\beta} 2^{2n}e^{-n} \geqq 2^{2n}e^{-n}.$$

Hence

$$R_j \leqq ac2^{-2n}e^n \leqq \frac{2(2n)!e^n}{n!n^n2^{2n}} < 2\sqrt{2},$$

using Stirling inequalities for the factorials.

To prove Theorem 4, suppose that H_4 is true. Then

$$P\left(w_n \geqq \frac{r}{n}\right) \leqq P\left(D(F_x^n,F_x) \geqq \frac{r}{n}\right) \leqq 2B(r,n) \leqq 4\sqrt{2}A(r,n),$$

and the proof of (2.8) proceeds as in Theorem 1. If H_4 is false and $\inf_{G\epsilon\mathcal{G}} D(F_x,G) = d > 0$, then for any G in $\mathcal{G}$,

$$D(F_x^n,G) \geqq D(F_x,G) - D(F_x^n,F_x) \geqq d - D(F_x^n,F_x),$$

so that $w_n \geqq d - D(F_x^n,F_x) \to d$ with probability 1 as $n \to \infty$ and hence $P(N < \infty) = 1$. The assertion concerning EN follows as before.

4. *Remarks.*—(*a*) It would be simple to extend Theorem 2 to test $H:D(F_x, F_y) \leqq c$ for any given $0 \leqq c < 1$.

(*b*) Some of the inequalities we use are quite crude, particularly the first inequality of (3.4), and the error probabilities are surely much smaller than we indicate.

(*c*) We could test H_3 by generating an artificial i.i.d. sequence $y_1,y_2,\ldots$ with d.f. F_0 and by applying Theorem 1. The procedure of Theorem 3, however, seems more natural.

(*d*) Possible choices of f satisfying conditions (*i*)–(*iv*) of Section 2 are given by the family

$$f(x) = ((x + 1)(a \log x + \log b))^{1/2}, \quad (a > 1, b \geqq 1) \tag{4.1}$$

for suitably large values of a, b, and m; e.g., if $x \geqq 1$ satisfies

$$\epsilon e^{x - 1} \geqq x + 1,$$

we can take $a = 2$, $b = 1/\epsilon(x^* - 1)$, $m = x^*$. In particular, for $\epsilon = 0.05$ we can take $a = 2$, $b = 4$, $m = 6$ in (4.1).

(*e*) The values of $f(x)$ for large x govern the behavior of EN for small d. More precisely, from (3.7) and (3.8) we have

$$1 \geqq \frac{y_0}{d} \geqq 1 - \frac{m}{f(g(d))} \to 1 \qquad \text{as } d \to 0,$$

so that $y_0 \sim d$ and $EN \leqq x_0 = g(y_0) \sim g(d)$ as $d \to 0$. If as in (4.1) $f(x)$

$c\sqrt{x \log x}$ as $x \to \infty$, it is simple to deduce that $g(x) \sim \frac{c}{x^2} \log \frac{1}{x}$ as $x \to 0$ and hence $EN = O\left(\frac{1}{d^2} \log \frac{1}{d}\right)$ as $d \to 0$.

It is interesting to compare this with the EN of a parametric test in reference 3. Suppose the x_i are $N(\mu,1)$, $\mu > 0$. Then $D(N(\mu,1),N(0,1)) \sim \frac{\mu}{\sqrt{2\pi}}$, $\mu \to 0+$, and hence, if H is false, the test of Theorem 3 applied to the hypothesis $H: \mu \leqq 0$ would have $EN = O\left(\frac{1}{\mu^2} \log \frac{1}{\mu}\right)$, $\mu \to 0+$. The test in reference 3, however, has $EN = O\left(\frac{1}{\mu^2} \log \log \frac{1}{\mu}\right)$, and this is the minimum possible order of magnitude.

(*f*) One cannot have $f(x) = O(\sqrt{x \log x})$, $x \to \infty$ if inequality (*iv*) is to hold. The *existence* of tests for which $f(x) = O(\sqrt{x \log \log x})$ follows from the results of Chung,[2] but we have not been able to exhibit them, owing to a lack of inequalities similar to those in the present paper.

(*g*) In addition to the bound (2.3) on EN, it can be shown that there exists $h > 0$ such that $E(e^{hN}) < \infty$.

* Research supported by NIH grant GM 10525-05 and by National Science Foundation grant GP-8381.

[1] Birnbaum, Z. W., and F. H. Tingey, "One-sided confidence contours for probability distribution functions," *Ann. Math. Statist.*, **22**, 592–596 (1951).

[2] Chung, K.-L., "An estimate concerning the Kolmogorov limit distribution," *Trans. Am. Math. Soc.*, **67**, 36–50 (1949).

[3] Darling, D. A., and H. Robbins, "Iterated logarithm inequalities," these PROCEEDINGS, **57**, 1188–1192 (1967). Cf. also *ibid.*, pp. 1577–1580; **58**, 66–68 (1967); and **60**, 1175–1182 (1968).

[4] Gnedenko, B. V., and V. S. Korolyuk, "On the maximal deviation between two empirical distributions," *Dokl. Akad. Nauk SSSR*, **80**, 525–528 (1951).

[5] Smirnov, N. V., "Approximation of random variables by empirical data," *Usp. Mat. Nauk*, **10**, 179–206 (1944).

[6] Whittaker, E. T., and G. S. Watson, *Modern Analysis* (London: Cambridge University Press, 1947).